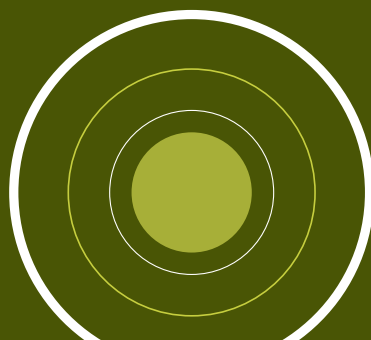


# Automatisierte Posteingangsbearbeitung Mailroom-Solutions Vergleichstest 2008/2009

›Wie leistungsfähig sind die führenden Posteingangsbearbeitungs-Produkte  
aus Sicht der Anwender und Experten?‹

**PENTADOC**  
Radar



Christoph Tylla, Analyst  
Guido Schmitz, Mitglied des Vorstandes  
Marc Weyhing, Analyst

© PENTADOC AG, Frankfurt am Main im November 2008

© Diese Studie ist urheberrechtlich geschützt. Alle Rechte vorbehalten.  
Der Nachdruck, auch auszugsweise, und die Wiedergabe als Ganzes oder in Auszügen unter Verwendung elektronischer Systeme ist nur mit schriftlicher Genehmigung des Autors gestattet.

Das vorliegende Dokument oder Teile davon dürfen nur für die Zwecke der Projektbearbeitung verwendet und an die relevanten Mitarbeiter verteilt werden. Das Dokument oder Teile davon dürfen weder im Original noch als Kopie an Dritte weitergegeben werden.

**Copyright**

Alle in diesem Dokument genannten (eingetragenen) Warenzeichen oder Schutzmarken stehen im Eigentum ihrer jeweiligen Besitzer bzw. Inhaber und werden von der PENTADOC AG anerkannt.

Bei der Verwendung von Zitaten und Abbildungen sind diese mit Quellenangaben zu versehen.

Gestaltung: ippolito fleitz group, Stuttgart – [www.ifgroup.org](http://www.ifgroup.org)

**1. Inhaltsverzeichnis**

<b>1. Inhaltsverzeichnis</b> .....	4
<b>2. Vorwort</b> .....	8
<b>3. Automatisierte Posteingangsbearbeitung Vergleichstest 2008/2009</b> .....	10
<b>4. Management Summary</b> .....	11
<b>5. Vergleichstest</b> .....	12
5.1 Motivation der PENTADOC und Pentadoc Radar zu einem Mailroom-Solutions-Vergleichstest .....	12
5.2 Auswahl der Produkte / Gipfeltreffen .....	13
5.3 Vorgehensweise Teststellung / Benchmark .....	14
5.3.1 Infrastruktur Testsystem .....	14
5.3.2 Virtualisierung .....	15
5.3.3 Durchführung der Benchmarks .....	16
5.3.4 Testdaten .....	17
5.3.4.1 Datenumfang .....	17
5.3.4.2 Dokumentenklassen .....	18
5.3.5 Abgleichdaten .....	19
5.4 Bewertungskriterien .....	21
5.4.1 »Vier Fahrzeuge des Premiumsegments im Vergleichstest« .....	21
5.4.2 Bewertungsmethodik im Detail .....	21
5.4.3 Erkennungsqualität .....	23
5.4.3.1 Messpunkte »Erkennungslauf Datenextraktion auf halbstrukturiertem Beleg« .....	24
5.4.3.2 Messpunkte »Klassifikation und Datenextraktion in unsortierten Posteingangsstapek« .....	25
5.4.4 Anwendungsbeispiele/Funktionen im Vergleich .....	26
5.4.5 Performance .....	26
5.4.6 Kriterienkatalog .....	26
5.4.7 Benutzerergonomie .....	28
<b>6. Automatisierte Posteingangsbearbeitung in der Theorie</b> .....	29
6.1 Definition ECM .....	29
6.2 Technologiekomponenten von ECM .....	30
6.2.1 Capture – Erfassung .....	30
6.2.2 Manage – Verwalten und Verarbeiten .....	31
6.2.3 Store – Speicherung .....	31
6.2.4 Preserve – Bewahrung .....	31
6.2.5 Delivery – Bereitstellung .....	31
6.3 Zukunftsweisende Technologiekomponenten im Bereich »Posteingangsbearbeitung« .....	32
6.4 Prozessablauf der elektronischen Posteingangsverarbeitung .....	36

6.5 Warum elektronischer Posteingang? .....	39
6.5.1 Rahmenbedingungen .....	39
6.5.2 Organisatorische Betrachtung .....	40
6.5.3 Was bringt es dem Anwender? .....	41
6.5.3.1 Zeitersparnis .....	42
6.5.3.2 Qualitätsverbesserung .....	42
6.5.3.3 Material- und Lagerkosten .....	43
6.5.4 Verfahrenserläuterung .....	43
6.5.5 Outsourcing der Posteingangsbearbeitung .....	45
6.5.6 Dokumenttypisierung: Dokumentenklassenmodell .....	45
6.6 Anbieter im Bereich Automatisierte Posteingangsbearbeitung .....	46
<b>7. Vorstellung der Teilnehmer</b> .....	<b>48</b>
7.1 EMC Deutschland GmbH .....	48
7.1.1 Unternehmensdaten .....	48
7.1.2.1 Produktinformationen .....	49
7.1.2.2 Produktbeschreibung .....	49
7.1.3 Komponenten und Screenshots .....	51
7.1.3.1 Manuelle Klassifizierung .....	51
7.1.3.2 Manuelle Feldkorrektur .....	52
7.1.3.3 Administrationsarbeitsplatz .....	53
7.1.4 Referenzen .....	54
7.1.5 Schulung, Beratung und technischer Support .....	55
7.1.6 Lizenzmodelle .....	55
7.1.7 Zukunftsausrichtung .....	56
7.1.8 Adresse, Ansprechpartner und weitere Informationen .....	56
7.2 Gaa GmbH & Co. KG (Gaa/smartFIX) .....	57
7.2.1 Unternehmensdaten .....	57
7.2.2 Produktdaten .....	59
7.2.2.1 Produktinformationen .....	59
7.2.2.2 Produktbeschreibung .....	59
7.2.3 Komponenten und Screenshots .....	60
7.2.3.1 Manuelle Klassifizierung .....	60
7.2.3.2 Manuelle Feldkorrektur .....	61
7.2.3.3 Administrationsarbeitsplatz .....	62
7.2.4 Referenzen .....	63

7.2.5 Schulung, Beratung und technischer Support .....	63
7.2.6 Lizenzmodelle .....	63
7.2.7 Zukunftsausrichtung .....	64
7.2.8 Adresse, Ansprechpartner und weitere Informationen .....	65
7.3 Kofax Deutschland AG .....	66
7.3.1 Unternehmensdaten .....	66
7.3.2 Produktdaten .....	67
7.3.2.1 Produktinformationen .....	67
7.3.2.2 Produktbeschreibung .....	67
7.3.3 Komponenten und Screenshots .....	69
7.3.3.1 Manuelle Klassifizierung .....	69
7.3.3.2 Manuelle Feldkorrektur .....	70
7.3.3.3 Administrationsarbeitsplatz .....	70
7.3.4 Referenzen .....	71
7.3.5 Schulung, Beratung und technischer Support .....	71
7.3.6 Lizenzmodelle .....	72
7.3.7 Zukunftsausrichtung .....	72
7.3.8 Adresse, Ansprechpartner und weitere Informationen .....	73
7.4 Top Image Systems (TIS) .....	74
7.4.1 Unternehmensdaten .....	74
7.4.2 Produktdaten .....	75
7.4.2.1 Produktinformationen .....	76
7.4.2.2 Produktbeschreibung .....	76
7.4.3 Komponenten und Screenshots .....	77
7.4.3.1 Manuelle Klassifizierung .....	77
7.4.3.2 Manuelle Feldkorrektur .....	78
7.4.3.3 Administrationsarbeitsplatz .....	79
7.4.4 Referenzen .....	80
7.4.5 Schulung, Beratung und technischer Support .....	81
7.4.6 Lizenzmodelle .....	81
7.4.7 Zukunftsausrichtung .....	82
7.4.8 Adresse, Ansprechpartner und weitere Informationen .....	83
<b>8. Die Systeme im Test .....</b>	<b>84</b>
8.1 Anwendungsbeispiele .....	84
8.2 Neuanlegen einer Anwendung – »time to application« .....	85

8.3 Funktionen im Vergleich .....	86
8.3.1 Parametrisierung der Klassifikation .....	87
8.3.2 Parametrisierung der Extraktion .....	94
8.3.3 Parametrisierung der Nachbearbeitungsmasken .....	101
8.3.4 Parametrisierung der unscharfen Suche .....	109
8.3.5 Workflow .....	110
8.3.6 Betriebssteuerung und Reporting .....	112
<b>9. Analyse und übergreifende Auswertung .....</b>	<b>117</b>
9.1 Ergebnisse des Benchmarks .....	117
9.2 Ergebnis Erkennungsqualität .....	117
9.2.1 Ergebnis Erkennungsqualität ›Erkennungslauf Datenextraktion auf halbstrukturiertem Beleg‹ .....	118
9.2.2 ›Erkennungslauf in unsortiertem Posteingangsstapel mit Datenbankabgleich‹ .....	120
9.2.2.1 Ergebnis Klassifikation .....	122
9.2.2.2 Ergebnis Datenextraktion auf Feldebene .....	123
9.2.2.3 Ergebnis Datenextraktion auf Dokumentebene .....	124
9.2.3 Ergebnis Performance .....	125
9.2.4 Ergebnis Benutzerergonomie .....	126
9.2.5 Ergebnis Kriterienkatalog .....	127
9.3 Gesamtergebnis .....	132
<b>10. Fazit .....</b>	<b>133</b>
<b>11. Dankeschön .....</b>	<b>135</b>
<b>12. Tabellen und Abbildungsverzeichnis .....</b>	<b>136</b>
12.1 Tabellenverzeichnis .....	136
12.2 Abbildungsverzeichnis .....	136

### 3. Automatisierte Posteingangsbearbeitung Vergleichstest 2008/2009

<b>Briefing</b>	Der Erstkontakt mit den Herstellern fand im Juli 2008 statt
<b>Zielergebnis</b>	Unter dem Schlagwort Automatisierte Posteingangsbearbeitung oder Mailroom-Solutions werden heute zahlreiche Lösungen angeboten. Doch was muss eine Mailroom-Solution können? Der Vergleichstest liefert Daten und Fakten zum Thema Automatisierungsgrad und berücksichtigt dabei auch wichtige Disziplinen wie Ergonomie und Performance.
<b>Zielgruppe</b>	Organisationsleiter, IT-Verantwortliche, Führungskräfte und ECM-Verantwortliche, die sich mit der Auswahl von Posteingangs- bzw. Mailroom-Solutions beschäftigen, aber auch die ECM-Hersteller zur Selbsteinschätzung.
<b>Teilnehmer</b>	EMC Deutschland GmbH Gaa GmbH/Insiders Technologies GmbH Kofax Deutschland AG Top Image Systems Deutschland GmbH
<b>Autoren</b>	Schmitz, Guido Tylla, Christoph Weyhing, Marc
<b>Auftraggeber</b>	PENTADOC AG Platz der Einheit 1, 60327 Frankfurt am Main Tel. +49 (0) 69 975 03 482 Fax +49 (0) 69 975 03 200 E-Mail: info@pentadoc.com
<b>Pentadoc</b>	Mit dem Geschäftsbereich Pentadoc Radar versorgt Pentadoc den deutschsprachigen ECM-Markt zukünftig mit detaillierten Marktdaten – Markt-Researches, Technologiestudien, Whitepapers, Benchmarks u.Ä. – und schließt damit die Informationslücke der großen Research-Häuser im breiten Spektrum des ECM-Marktes.

#### 4. Management Summary

Gerade in wirtschaftlich schwierigen Zeiten streben Unternehmen einen immer höheren Automatisierungsgrad an, um Kosten zu sparen und wettbewerbsfähig zu bleiben. Im Hinblick auf Prozesseffizienz, aber auch um den steigenden Anforderungen des Gesetzgebers gerecht zu werden, ist ein logischer Schritt, den bestehenden Posteingang elektronisch zu bearbeiten. Dies ist der Übergang von einer papiergestützten Verarbeitung und Archivierung der eingehenden Postdokumente zu einer elektronischen – sprich eine Automatisierung der Posteingangsbearbeitung.

Doch leider kann die übliche Vorgehensweise einer Produktauswahl – Erarbeitung eines Pflichtenhefts, Kriterienkatalog, Angebotsvergleich und ggf. noch Anbieterpräsentation – selten den tatsächlichen Leistungsumfang oder die Leistungsgrenze der betrachteten Lösungen darlegen. Theoretisch funktionieren alle am Markt angebotenen Produkte sehr gut, aber tun sie dies auch mit dem realen Beleggut? Oder nur mit den Beispielen der Verkaufspräsentation? Zudem fehlen in den im Allgemeinen sehr engen Projektplänen einfach die Wochen oder Monate, in welchen die Systeme vor Ort auf Herz und Nieren mit eigenen Belegen getestet werden können. Und neben dieser fehlenden Zeit ist auch das Know-how, die Tests vorzubereiten, durchzuführen und auszuwerten, nur in den seltensten Fällen in den Organisationen verfügbar.

Um diesem Umstand Rechnung zu tragen, hat die PENTADOC zu einem Benchmark aufgerufen, in dem die aktuell führenden Systeme gegenübergestellt werden sollten. Als Vorgehensweise wurde die eines Labortests gewählt. Die Systeme wurden auf derselben Hardware-Umgebung installiert und mit den gleichen Dokumenten – dem Posteingang – gefüttert. Die Ergebnisse können also repräsentativ miteinander verglichen werden.

Mit dem Vergleichstest sollte herausgefunden werden, welches angebotene System mit den Anforderungen eines gemischten, heterogenen Posteingangs am besten zurechtkommt. Neben dieser mathematisch sehr einfach mittels Zählen von gut und schlecht erkannt bzw. klassifizierten Dokumenten zu bewertenden Eigenschaft wurden jedoch noch die »klassischen« Ausschreibungspunkte betrachtet. Es wurde also auch eigens für diesen Testlauf ein Pflichtenheft erstellt und eine Art Präsentation der Produkte eingefordert, um diese auszuwerten.

Um die ursprüngliche Idee nicht aus den Augen zu verlieren, nimmt die eigentliche Leistung der Systeme den größten Stellenwert in der Bewertung ein. Wie schon erwähnt, konnten hier mathematisch objektiv die Ergebnisse gefunden werden. Diese Ergebnisse können Sie im Kapitel 9 für die einzelnen Testdisziplinen nachlesen. Ein kleiner Vorgeschmack auf die Ergebnisse: Wenn man so wie in diesem Benchmark mit einem Lineal objektiv misst, gibt es natürlich meistens einen »Klassenbesten« – so auch hier. Als Gesamtsieger der Teststellung geht die Firma TIS mit Ihrem Produkt eFlow hervor.

Allerdings gibt es eigentlich keinen Verlierer, da alle Systeme ihre Leistungsfähigkeit unter Beweis gestellt haben und jeweils als führende Produkte in diesem Marktsegment rangieren. Und daher sei an dieser Stelle auch schon sehr frühzeitig eine erste Warnung ausgesprochen: So repräsentativ die Ergebnisse im Rahmen dieses Benchmarks sein mögen, sie sind es ganz bestimmt nicht für jede Unternehmung bzw. für jedes Projekt! Erst wenn man die ganz genauen Anforderungen und Rahmenbedingungen – wie Zusammenspiel mit Fachanwendungen und aktueller Systemarchitektur – kennt, kann man für den jeweiligen Zweck das richtige Produkt auswählen.

Eine Studie und die daraus resultierenden Bewegungen bei Kunden und Herstellern leben auch von Ihrer ganz persönlichen Meinung. Teilen Sie uns bitte mit, welche Erfahrungen Sie selbst gemacht haben, welche Aspekte besonders gut beleuchtet wurden oder wo Sie noch Nachholbedarf sehen. Gerne berücksichtigen wir Ihre Anmerkungen in unseren nächsten Tests und können Ihnen dann so noch nützlichere Informationen bieten.

## 5.2 Auswahl der Produkte / Gipfeltreffen

Wie bereits der Titel der zweiten Auflage des Vergleichstests >Gipfeltreffen Automatisierte Posteingangsbearbeitung / Mailroom-Solutions< suggeriert, galt es die führenden Lösungen in diesem Segment zu finden. Zielsetzung war, aus über zwanzig am Markt verfügbaren Lösungen die führenden in diesem Marktsegment zu ermitteln und die Ergebnisse des Vergleichstests als >Richtschnur< für weitere Produktevaluierungen heranzuziehen.

Was heißt aber führende Lösung im Bereich Automatisierte Posteingangsbearbeitung? Es spricht eine Reihe von Faktoren für eine Marktführerschaft. In dem vorliegenden Vergleichstest wurde aber im speziellen auf die Faktoren Funktionsumfang, Technologische Plattform, Marktpräsenz, Referenzkunden und Unternehmen Wert gelegt. Vor der eigentlichen Teststellung wurden daher durch die Analysten der Pentadoc Radar die Lösungen der am Markt befindlichen Hersteller bewertet und in eine Short-List neun Lösungen überführt.

Voraussetzung für die folgende Teilnahme am Vergleichstest war, dass die jeweilige Lösung kostenfrei zur Verfügung zu stellen und mindestens ein Produktfachmann zur Installation und Einweisung in das System und zum Aufbau der Anwendungsbeispiele für zwei Tage abzustellen war.

Die aktuellen Entwicklungen und der komplette Umfang der gezeigten Lösung, neben allgemeinen Fragen zum Lösungs

### 6.3 Zukunftsweisende Technologiekomponenten im Bereich ›Posteingangsbearbeitung‹

Im Bereich Automatisierte Posteingangsbearbeitung sind zwei Technologiekomponenten von besonderem Interesse: Die Klassifikation und die Extraktion.

Um zu verstehen wie diese Komponenten arbeiten und welche Potenziale aufgrund neuer wissenschaftlicher Methoden noch möglich sind, haben wir einen Gastartikel von Dipl. Inform. Heiko Groftschik in diesem Kapitel eingefügt. Herr Groftschik ist Spezialist für computerlinguistische Forschung und leitet die Entwicklung beim Software-Lieferanten ITyX. Die entwickelten Lösungen kommen bei der Eingangsklassifikation und -verarbeitung von digitalen Geschäftsvorfällen per E-Mail, Brief oder Fax (Enterprise Response Management) oder bei der Informationsbewertung in mittleren bis großen Organisationen (Enterprise Content Management) zum Einsatz. [www.ityx.de](http://www.ityx.de)

#### ›Die vollautomatische Erkennung und Verarbeitung unstrukturierter Texte mittels wissenschaftlicher Methoden optimiert den Wirkungsgrad klassischer ECM- und Posteingangs-Lösungen.‹

Dipl. Inform. Heiko Groftschik

Eine neue Generation von innovativen Software-Lösungen ist heute in der Lage, Textinformationen unabhängig von Struktur oder Sprache automatisch zu kategorisieren. Zum Einsatz kommen komplexe wissenschaftliche Algorithmen, deren Verwendung erstaunliche Potenziale für die Informationsbewertung und -verarbeitung eröffnet. Optimal zum Einsatz gebracht können diese Klassifikationsverfahren den Automatisierungsgrad in mittleren bis großen Serviceorganisationen erheblich steigern. Dabei ist es unerheblich, ob die Verfahren bei der Verarbeitung von digitalen Eingangspoststücken (E-Mail, Brief- und Faxverarbeitung) zum Einsatz kommen oder zur Suche und Bewertung von Informationen innerhalb der Organisation (Enterprise Content Management).

Im Zeitalter des Internets ist die Kommunikation zwischen Verbrauchern und serviceorientierten Organisationen digital und schnell geworden. Alles dreht sich um den Faktor Zeit. Jede eingesparte Sekunde bei der Verarbeitung einzelner Geschäftsvorfälle multipliziert sich in Arbeitsabläufen von Konzernen in unvorstellbare Summen, die möglichst auf der Haben-Seite verbucht und nicht auf der Soll-Seite abgeschrieben werden wollen. Insbesondere in der schriftbasierten Kommunikation mit Kunden und Interessenten, der schnellen Verfügbarkeit vorgangsbasierter Daten und der Archivierung von Dokumenten und Informationen bietet der Einsatz von wissenschaftlichen Methoden der Text- und Informationsverarbeitung ein enormes Potenzial für Prozessautomatisierungen und Rentabilität.

In der Praxis werden insbesondere teil- und unstrukturierte Kundenmitteilungen in Serviceorganisationen gleich durch mehrere Mitarbeiter gelesen und bewertet. Die durchschnittlichen Bearbeitungszeiten je Geschäftsvorfall sind hoch. Zunächst wird festgestellt, wer für ein Problem zuständig ist. Erst dann wird der Vorgang einer zuständigen Person oder Arbeitsgruppe zugestellt, nochmals erfasst und zur Auslösung weiterführender Verarbeitungsschritte weitergeleitet. Insbesondere bei der Verarbeitung digitaler Mitteilungen per E-Mail an die sogenannten ›info(Q‹-Postkörbe oder eingehender Briefpost an zentrale Firmenadressen fällt eine mögliche Vorauswahl einer zuständigen Abteilung vollständig aus. Hier kann schon mit klassischen Kategorisierungslösungen wie beispielsweise der Verwendung von Schlagwörtern, Namenslisten von Sachbearbeitern und regulären Ausdrücken eine gewisse Vorauswahl getroffen werden. Diese Methoden sind offensichtlich nur rudimentär verwendbar, besitzen sie doch meist eine hohe False-Positive-Rate:<sup>1</sup> Dokumente werden Kategorien oder Mitarbeitern falsch zugeordnet und führen zu einem hohen manuellen Arbeitsaufwand, der den Weg für zahllose Fehlerquellen und Missverständnisse öffnet.

<sup>1</sup> Maß zur Beurteilung der Güte von Treffermengen einer Recherche beim Information Retrieval

## 8. Die Systeme im Test

### 8.1 Anwendungsbeispiele

Für den Vergleichstest wurden zwei reale Szenarien als Drehbücher und Anhaltspunkte für die Teilnehmer zur Vorstellung und anschließenden Bewertung der Posteingangslösungen durch die Prüfer der Pentadoc Radar verwendet. Im Konkreten waren dies:

1. Neuanlegen einer Anwendung zur Klassifikation und Extraktion am Beispiel der Dokumenttypen Adressänderung und Kündigung,
2. Optimierung und Betrieb der Anwendung.

Bei den Anwendungsbeispielen wurde insbesondere die ›time to application‹ bewertet. Unter ›time to application‹ wird der Zeitraum verstanden, der von einem geschulten Anwendungsentwickler benötigt wird, um eine funktionierende Anwendung zu erstellen. Neben der Zeit war natürlich bei den Anwendungsszenarien auch relevant, welches Know-how bei den vorgestellten Lösungen notwendig war, um eine neue Anwendung anzulegen und diese dann zu optimieren bzw. im Betrieb zu überwachen.

Insgesamt wurden im Bereich der Anwendungsbeispiele 10 Punkte für die ›time to application‹ vergeben und 10 Punkte für die gezeigten Bearbeitungsschritte, die Funktionen im Vergleich und das dazu notwendige Entwicklungs-Know-how auf Anwenderseite.

Die Parameter ›time to application‹ und ›Optimierung und Betrieb von Applikationen‹ sind auch neben den Investitionskosten die maßgeblichen Faktoren bei der Berechnung des Total Cost of Ownership (TCO) im Bereich Automatisierte Posteingangsbearbeitung.

Unter dem TCO versteht man im Allgemeinen den Ansatz, Unternehmen dabei zu helfen, alle anfallenden Kosten von Investitionsgütern in der IT wie beispielsweise Software abzuschätzen. Die Idee dabei ist, eine Abrechnung zu erhalten, die nicht nur die Anschaffungskosten enthält, sondern alle Aspekte der späteren Nutzung der betreffenden Komponenten. Somit können bekannte Kostentreiber oder auch versteckte Kosten möglicherweise bereits im Vorfeld einer Investitionsentscheidung identifiziert werden.

## 8.2 Neuanlegen einer Anwendung – >time to application<

Am Beispiel der Dokumenttypen Adressänderung und Kündigung wurde mit entsprechenden Trainingsdokumenten der Aufbau einer neuen Anwendung für die Verarbeitung von gemischtem Beleggut gezeigt.

Dabei wurden insbesondere folgende Funktionen geprüft:

- Parametrisierung der Image-Vorverarbeitung (inklusive Leerseitenerkennung)
- Parametrisierung der Dokumenttrennung (Patchcode-Blatt, Barcode, ...)
- Parametrisierung der Klassifikation
- Parametrisierung der Extraktion an den Feldern Versicherungsnummer
- Parametrisierung der Nachbearbeitungsmasken
- Parametrisierung der unscharfen Suche

Für die Erreichung der maximalen Punktzahl in diesem Bereich wurde eine Zeit von vier Stunden angenommen, d. h. konnten beide Applikationen in 240 Minuten angelegt werden, so wurde die maximale Punktzahl von 10 vergeben. Im Gegenzug wurde als Obergrenze eine Zeitdauer von 8 Stunden festgelegt, d. h. konnten innerhalb von 480 Minuten die beiden Applikationen nicht angelegt werden, so wurde dieser Bereich mit 0 Punkten bewertet.

Insgesamt konnten alle Teilnehmer das Neuanlegen von einer Anwendung in der vergebenen Zeit absolvieren.

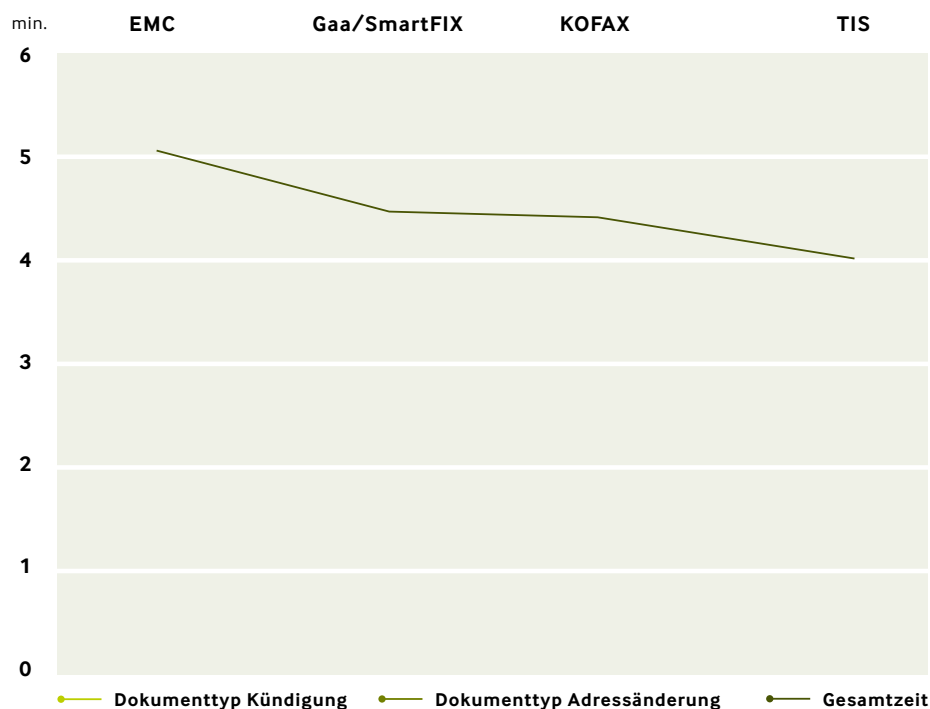


Abbildung 19: Auswertung >time to application<



Gaa/smartFIX verwendet zur Extraktion schwerpunktmäßig den Erkennungsalgorithmus der Firma Kadmos, welcher das gelesene Ergebnis in einem sogenannten Hypothesengraph ausgibt. Der Hypothesengraph schlägt je gelesenem Zeichen mehrere Ergebniszeichen vor. So kann mit jedem Vorschlagszeichen z.B. der Datenbankabgleich gestartet werden, bis die richtige Übereinstimmung gefunden wurde.

TIS und EMC binden als Unterstützung zur Verbesserung der Extraktionsergebnisse die Datenbanksuche der Firma Exorbyte ein. Mit diesem Ansatz ist es möglich, den unscharfen Datenbankabgleich unternehmensweit auch für andere Anwendungen zu verwenden, wo hingegen die proprietären Datenbankabgleiche von Gaa/smartFIX und Kofax dahingehend einen Nachteil haben: sie können nur in der Mailroom-Solution eingesetzt werden. Alle Hersteller lassen ein hohes Maß dialogorientierter Parametrisierung durch den Anwender zu.

### Die Ergebnisse zum Kapitel im Überblick

Anbieter	EMC	Gaa/smartFIX	Kofax	TIS
Bewertung	+	-	-	+

### Welche Punkte sind in diesem Funktionsbereich besonders aufgefallen?

#### Das gefällt den Anwendern

Alle Firmen bieten umfangreiche Optimierungsmöglichkeiten an.

Alle Teilnehmer bieten die Möglichkeit Datenbanken zur Verbesserung der Leseergebnisse einzubinden.

Den flexibelsten Eindruck macht die Firma TIS aufgrund der großen Anzahl der integrierten Erkennungsesines.

Alle Systeme bieten eine durchgängige Oberfläche für eine benutzerfreundliche Parametrisierung der Extraktion.

Alle Produkte bieten ein Voting an. Die meisten Parametrisierungsmöglichkeiten im Voting bietet die Firma TIS.

#### Das gefällt den Anwendern nicht

Proprietäre in das Produkt gekapselte Datenbankabgleiche von Gaa/smartFIX und Kofax.

Die Definition von Lokatoren bei Kofax erscheint komplex und skriptlastig.

Kein dezidierter formularbasierender Ansatz der Firma Kofax.

## 9. Analyse und übergreifende Auswertung

### 9.1 Ergebnisse des Benchmarks

Die Gesamtbewertung des Posteingangsbenchmarks setzt sich zusammen aus insgesamt fünf Bereichen, welche unterschiedlich gewichtet wurden.

1. Erkennungsqualität
2. Anwendungsbeispiele/Funktionen im Vergleich
3. Performance
4. Kriterienkatalog
5. Benutzeroberflächen

Die detaillierte Beschreibung der einzelnen Kriterien folgt in den nachfolgenden Kapiteln.

### 9.2 Ergebnis Erkennungsqualität

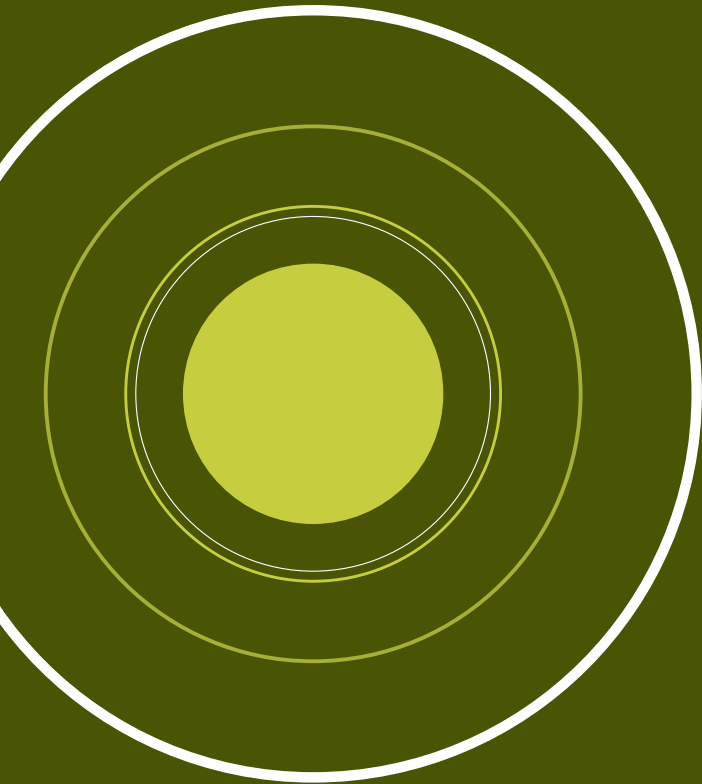
Wie bereits im Kapitel 5.4.3 erläutert, unterteilt sich das Ergebnis der Erkennungsqualität in die Bereiche ›Erkennungslauf Datenextraktion auf halbstrukturiertem Beleg‹ und ›Erkennungslauf in unsortiertem Posteingangsstapel mit Datenbankabgleich‹.

In den beiden Messungen wurde dann jeweils die Erkennungsqualität in den Bereichen ›Klassifikation‹ und ›Datenextraktion‹ ermittelt. Die Datenextraktion selbst ist nochmals untergliedert in die Bereiche ›Zeichenebene‹, ›Feldebene‹ und ›Dokumentebene‹.

<b>EMC</b>	<b>58,75 %</b>
<b>Gaa/smartFIX</b>	<b>72,89 %</b>
<b>Kofax</b>	<b>74,42 %</b>
<b>TIS</b>	<b>73,66 %</b>

Abbildung 59: Ergebnisdiagramm Erkennungsqualität

Im Bereich der Erkennungsqualität konnten die Firmen Kofax, TIS und Gaa/smartFIX überzeugen. Die von den drei Firmen gezeigten Klassifikations- und Extraktionsergebnisse können für weitere Vergleichstests mit vergleichbarer Aufgabenstellung als Richtwert des maximal Erreichbaren mit dem aktuellen Stand der Technik verwendet werden. Insgesamt konnte die Firma Kofax im Bereich der Erkennungsqualität mit dünnem Vorsprung das Rennen machen.



**PENTADOC RADAR** versorgt den deutschsprachigen ECM-Markt zukünftig noch fokussierter mit detaillierten Marktdaten – Business-Researches, Technologiestudien, White Papers, Benchmarks u. Ä. – und schließt damit die Informationslücke der großen Research-Häuser im breiten Spektrum des ECM-Marktes.

PENTADOC AG  
Kastor & Pollux  
Platz der Einheit 1 D-60327  
Frankfurt  
Fon +49 (0) 69 975 03 482  
Fax +49 (0) 69 975 03 200  
Mail [info@pentadoc.de](mailto:info@pentadoc.de)

PENTADOC Schweiz GmbH  
Baarerstraße 63 CH-6302 Zug  
Fon +41 (0) 41 729 09 62  
Fax +41 (0) 41 729 09 99  
Mail [info@pentadoc.ch](mailto:info@pentadoc.ch)

PENTADOC Ges. m.b.H.  
Börsegebäude  
Schottenring 16 A-1010 Wien  
Fon +43 (0) 1 537 124 818  
Fax +43 (0) 1 537 124 000  
Mail [info@pentadoc.at](mailto:info@pentadoc.at)

[www.pentadoc.com/radar](http://www.pentadoc.com/radar)